Nonlinear denoising of transient signals with application to event related potentials

A. Effern [a,b], K. Lehnertz [a], T. Schreiber [c], T. Grunwald [a],
P. David [b], C.E. Elger [a]

[a]*Department of Epileptology, University of Bonn, Sigmund-Freud Str. 25,
53105 Bonn, Germany*

[b]*Institute of Radiation and Nuclear Physics, University of Bonn, Nussallee 11-13,
53115 Bonn, Germany*

[c]*Department of Physics, University of Wuppertal, Gauss-Strasse 20,
42097 Wuppertal*

Abstract

We present a new wavelet based method for the denoising of *event related potentials* (ERPs), employing techniques recently developed for the paradigm of deterministic chaotic systems. The denoising scheme has been constructed to be appropriate for short and transient time sequences using circular state space embedding. Its effectiveness was successfully tested on simulated signals as well as on ERPs recorded from within a human brain. The method enables the study of individual ERPs against strong ongoing brain electrical activity.

Keywords: nonlinear denoising, state space, wavelets, circular embedding

1 Introduction

The *electroencephalogram* (EEG) reflects brain electrical activity owing to both intrinsic dynamics and responses to external stimuli. To examine pathways and time courses of information processing under specific conditions, several experiments have been developed controlling sensory inputs. Usually, well defined stimuli are repeatedly presented during experimental sessions (e.g., simple tones, flashes, smells, or touches). Each stimulus is assumed to induce synchronized neural activity in specific regions of the brain, occurring as potential changes in the EEG. These *evoked potentials* (EPs) often exhibit multiphasic peak amplitudes within the first hundred milliseconds after stimulus onset. They are specific for different stages of information processing, thus giving access to both temporal and spatial aspects of neural processes. Other classes of experimental setups are used to investigate higher cognitive functions. For example, subjects are requested to remember words, or perhaps they are asked to respond to specific target stimuli, e.g. by pressing a button upon their occurrence. The neural activity induced by this kind of stimulation also leads to potential changes in the EEG. These *event related potentials* (ERPs) can extend over a few seconds, exhibiting peak amplitudes mostly later than EPs. Deviation of amplitudes and/or moment of occurrence (latency) from those of normal EPs/ERPs are often associated with dysfunction of the central nervous system and thus, are of high relevance for diagnostic purposes.

As compared to the ongoing EEG, EPs and ERPs possess very low peak amplitudes which, in most cases, are not recognizable by visual inspection. Thus, to improve their low signal-to-noise ratio, EPs/ERPs are commonly averaged (Figure 1), assuming synchronous, time-locked responses not correlated with the ongoing EEG. In practice, however, these assumptions may be inaccurate

and, as a result of averaging, variations of EP/ERP latencies and amplitudes are not accessed. In particular, short lasting alterations which may provide relevant information about cognitive functions are probably smoothed or even masked by the averaging process. Therefore, investigators are interested in *single trial analysis*, that allows extraction of reliable signal characteristics out of single EP/ERP sequences [1]. In ref. [2] **auto**regressive models (AR) are adopted to EEG sequences recorded prior to stimulation in order to subtract uncorrelated neural activity from ERPs. However, it is an empirical fact, that external stimuli lead to event-related-desynchronizaition of the ongoing EEG. Thus, the estimated AR-model might be incorrect. The authors of [3] applied **auto**regressive **m**oving **a**verage ($ARMA$) models to time sequences which were a concatenation of several EP/ERP sequences. In the case of short signal sequences, this led to better spectral estimations than commonly achieved by periodograms. The main restriction is, however, that investigated signals must be linear and stationary, which cannot be strictly presumed for the EEG. In particular the high model order in comparison to the signal length shows that AR- and ARMA-models are often inadequate for EP/ERP analysis. Other methods have been developed to deal with the nonstationary and transient character of EPs/ERPs. Woody [4] introduced an iterative method for EP/ERP latency estimation based on common averages. He determined the time instant of the best correlation between a template (EP/ERP average) and single trials by shifting the latter in time. This method corrects a possible latency variability of EPs/ERPs, but its performance highly depends on the initial choice of templates. The *Wiener filter* [5,6], on the other hand, uses spectral estimation to reduce uncorrelated noise. This technique, however, is less accurate for EPs/ERPs, because the time course of transient signals is lost in the Fourier domain. Thus, DeWeerd [7,8] introduced a time adaptive Wiener filter, allowing better adjustment to signal components of short du-

ration. The paradigm of orthogonal wave packets (*wavelet transform*[1]) also follows this concept of adopted time-frequency decomposition. In addition, the wavelet transform provide several useful properties which make it preferable even for the analysis of transient signals [9–11]:

- Wavelets can represent smooth functions as well as singularities.
- The basis functions are local which makes most coefficient based algorithms to be naturally adapted to inhomogeneities in the function.
- They have the unconditional basis property to represent a variety of functions implying that the wavelet basis is usually a reasonable choice even if very little is known about the signal.
- Fast wavelet transform is computationally inexpensive of order $O(N)$, where N denotes the number of sample points. In contrast, fast Fourier transform (FFT) requires $O(Nlog(N))$.
- Nonlinear thresholding is nearly optimal for signal recovery.

For that reasons, wavelets became a popular tool for the analysis of brain electrical activity [12–15], especially for denoising and classification of single trial EPs/ERPs. Donoho et al.[16] introduced a simple thresholding algorithm to reduce noise in the wavelet domain requiring no assumptions about the time course of signals. Nevertheless, high signal amplitudes are in need to distinguish between noise and signal related wavelet coefficients in single trials. Bertrand et al. [17] modified the original *a posteriori Wiener filter* to find accurate filter settings. The authors emphasized better adoption to transient signal components than can be achieved by corresponding techniques in the frequency domain. However, due to the averaging process, this technique runs

---

[1] Continuous wavelet transform: $w_{a,b}(\Psi, x(t)) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t)\Psi(\frac{t-b}{a})dt$

w: wavelet coefficient, a: scaling parameter, b: translation parameter, $x(t)$: time series, Ψ: *mother wavelet* function

the risk of choosing inadequate filter settings in the case of a high latency variability. The same restriction is valid for discriminant techniques applied e.g. by Bartink et al. [18,19]. Nevertheless, wavelet based methods enable a more adequate treatment of transient signals than techniques applied in the frequency domain. The question of accurate filter settings, however, is still an unresolved problem.

To circumvent this problem, we introduce a new method for single trial analysis of ERPs that neither assumes fully synchronized nor stationary ERP sequences. The method is related to techniques already developed for the paradigm of deterministic chaotic systems, using time delay embeddings of signals for state space reconstruction and denoising [20]. Schreiber and Kaplan [21] demonstrated the accuracy of these methods to reduce measurement noise in the human electrocardiogram (ECG). Heart beats are also of transient character and exhibit relevant signal components in a frequency range that compares to ERPs. Unfortunately, ERPs are of shorter duration as compared to the ECG. Thus, in the case of high dimensional time delay embedding (in the order of the signal length), we cannot create a sufficient number of delay vectors for ERP sequences. To circumvent this problem we reconstruct ERPs in state-space using circular embeddings, that have turned out to be appropriate even for signal sequences of short duration. In contrast to the nonlinear projection scheme described in [20], we do not use *singular value decomposition* (SVD) to determine clean signals in state space. The reason for this is threefold. First, estimating relevant signal components using the inflexion of ordered eigen-values is not always applicable to EEG because eigen-values may decay almost linearly. In this case, an a priori restriction to a fixed embedding dimension is in need, running the risk either to discard important signal components or to remain noise of considerable amplitude if only little is known

about the signal. Second, SVD stresses the direction of highest variances, so that transient signal components may be smoothed by projection. Third, the number of signal related directions in state space may alter locally, which is also not concerned by SVD. Instead we calculate wavelet transforms of delay vectors and determine signal related components by estimating variances separately for each state-space direction. Scaling properties of wavelet bases allow very fast calculation as well as focusing on specific frequency bands. To confirm the accuracy of our method, we apply it to ERP-like test signals contaminated with different types of noise. Afterwards, we give an example of reconstructed mesial temporal lobe P300 potentials, that were recorded from within the hippocampal formation of a patient with focal epilepsy.

2 Outline of the Method

A time series may be contaminated by random noise allowing the measurement $y_n = x_n + \epsilon_n$. If the measured time series is purely deterministic, it is restricted to a low-dimensional hyper-surface in state space. For the transient signals we are concerned with here, we assume this still to be valid. We hope to identify this direction and to correct y_n by simply projecting it onto the subspace spanned by the clean data [22,21].

Technically we realize projections onto noise free subspaces as follows. Let $Y = (y_1, y_2, \ldots, y_N)$ denote an observed time sequence. Time-delay embedding of this sequence in a m-dimensional state space leads to state space vectors $\mathbf{y}_n = (y_n, \ldots, y_{n-(m-1)\tau})$, where τ is an appropriate time delay. In an embedding space of dimension m we compute the discrete wavelet transform [11,10,9] of all delay vectors in a small neighborhood of a vector $\mathbf{y}_n$ we want to correct. Let $r_{n,j}$ with $j = 0, \ldots, k$ denote the indices of the k nearest neighbors of $\mathbf{y}_n$, and

for $\mathbf{y}_n$ itself, i.e. $j = 0$, and $r_{n,0} = n$. Thus, the first neighbor distances from $\mathbf{y}_n$ in increasing order are $d(Y)_n^{(1)} \equiv ||\mathbf{y}_n - \mathbf{y}_{r_{n,1}}|| = \min_{r'} ||\mathbf{y}_n - \mathbf{y}_{r'}||$, $d(Y)_n^{(2)} \equiv ||\mathbf{y}_n - \mathbf{y}_{r_{n,2}}|| = \min_{r' \neq r_{n,1}} ||\mathbf{y}_n - \mathbf{y}_{r'}||$, etc., where $||\mathbf{y} - \mathbf{y}'||$ is the Euclidean distance in state space. Now the important assumption is that the clean signal lies within a subspace of dimension $d \ll m$, and that this subspace is spanned by only a few basis functions in the wavelet domain. Let $\mathbf{w}_{r_{n,j}}$ denote the fast wavelet transform [23,24] of $\mathbf{y}_{r_{n,j}}$. Futhermore, let $C_i^{(k)}(\mathbf{w}_{r_n}) = \langle \mathbf{w}_{r_{n,j}} \rangle_i$ denote the ith component of the centre of mass of $\mathbf{w}_{r_n}$, and $\sigma_{n,i}^2$ the corresponding variance. In the case of neighbors owing to the signal (*true neighbors*), we can expect the ratio $C_i^{(k)}(\mathbf{w}_{r_n})/\sigma_{n,i}^2$ to be higher in signal than in noise related directions. Thus, a discrimination of noise and noise free components in state space is possible. Let

$$\tilde{w}_{n,i} = \begin{cases} w_{n,i} & : & |C_i^{(k)}(\mathbf{w}_{r_n})| \geq 2\lambda \frac{\sigma_{n,i}}{\sqrt{k+1}} \\ 0 & : & \text{else} \end{cases} \tag{1}$$

define a shrinking condition to carry out projection onto a noise free manifold [16]. The parameter λ denotes a thresholding coefficient that depends on specific qualities of signal and noise. Inverse fast wavelet transform of $\tilde{\mathbf{w}}_n$ provides a corrected vector in state space, so that application of our projection scheme to all remaining delay vectors ends up with a set of corrected vectors, out of which the clean signal can be reconstructed.

2.1 Extension to multiple signals of short length

Let $Y_l = (y_{l,1}, y_{l,2}, \ldots, y_{l,N})$ denote a short signal sequence that is repeatedly recorded during an experiment, where $l = 1, \ldots, L$ orders the number of repetitions. A typical example may be ERP recordings, where each Y_l represents an EEG sequence following well defined stimuli. Time-delay embeddings of

these sequences can be written as $\mathbf{y}_{l,n} = (y_{l,n} \ldots, y_{l,n-(m-1)\tau})$. To achieve a sufficient number of delay vectors even for high embedding dimensions, we define circular embeddings by

$$\mathbf{y}_{l,n} = (y_{l,n}, \ldots, y_{l,1}, y_{l,N}, \ldots, y_{l,N-(m-q)}) \quad \forall \quad n < m, \tag{2}$$

so that all delay vectors with indices $1 \leq n \leq N$ can be formed. Circular embeddings are introduced as the most attractive choice to handle the ends of sequences. Alternatives are (i) losing neighbors, (ii) zeropadding, and (iii) shrinking the embedding dimension towards the ends. However, discontinuities may occur at the edges, requiring some smoothing. For each Y_l we define the smoothed sequence as

$$\mathbf{y}_{l,n,i}^{s} = \begin{cases} y_{l,n,i} e^{-(\frac{q-i}{p})^2} & : \quad i < q \\ y_{l,n,i} & : \quad q \leq i \leq N - q \\ y_{l,n,i} e^{-(\frac{i-(N-q)}{p})^2} & : \quad i > N - q \end{cases} \tag{3}$$

where q defines the window width in sample points, p the steepness of exponential damping, and i the time index. Time-delay embedding of several short sequences leads to a filling of the state space, so that a sufficient number of nearest neighbors can be found for each point.

2.2 Parameter Selection

Appropriate choice of parameters, in particular embedding dimension m, time delay τ, thresholding coefficient λ, as well as the number of neighbors k is important for accurate signal reconstruction in state space. Several methods have been developed to estimate "optimal" parameters, depending on specific aspects of the given data (e.g., noise level, type of noise, stationarity, etc.). These assume that the clean signal is indeed low dimensional, an assumption

we are not ready to make in the case of ERPs. Thus, we approached the problem of "optimal" parameters empirically.

Parameters τ and m are not independent from each other. In particular, high embedding dimensions allow small time-delays and vice versa. We estimated "optimal" embedding dimensions and thresholding coefficients on simulated data by varying m and λ for a fixed $\tau = 1$. To allow fast wavelet transform, we chose m to be a power of 2.

Repeated measurements, like in the case of EPs/ERPs, have a maximum number of true neighbors which is given by $k_{max} = L$. In the case of identical signals this is the best choice imaginable. However, real EPs/ERPs may alter during experiments, and it seems more appropriate to use a maximum distance true neighbors are assumed to be restricted to. We define this distance by

$$ d(\mathbf{y})_{max} = \frac{\sqrt{2}}{LN} \sum_{l=1,n=1}^{L,N} d(\mathbf{y})_{n,l}^{(L)} \tag{4} $$

3   Model Data

3.1   Generating test signals and noise

To demonstrate the effectiveness of our denoising technique and to estimate accurate values for m, λ, and L, we applied it to EP/ERP-like test signals contaminated with white noise and in-band noise. The latter was generated using phase randomized surrogates of the original signal [25]. Test signals consisted of 256 sample points and were a concatenation of several Gaussian functions with different standard deviations and amplitudes. To simulate EPs/ERPs not fully synchronized with stimulus onset, test signals were shifted randomly in time (normally deviated, std. dev.: 20 sample points, max. shift: 40 sample

points). Since even fast components of the test signal extended over several sample points, a minimum embedding dimensions $m = 16$ was required to cover any significant fraction of the signal. The highest embedding dimension was bounded by the length of signal sequences and the number of embedded trials, thus allowing a maximum of $m = 256$. However, if the embedding dimension is $m = N$, neighborhood is not longer defined by local characteristics, and we can expect denoised signals to be smoothed in the case of multiple time varying components.

3.2 Denoising of test signals

Let $X_l = (x_{l,1}, x_{l,2}, \ldots, x_{l,N})$ denote the lth signal sequence of a repeated measurement, $Y_l = (y_{l,1}, y_{l,2}, \ldots, y_{l,N})$ the noise contaminated sequence, and $\tilde{Y}_l = (\tilde{y}_{l,1}, \tilde{y}_{l,2}, \ldots, \tilde{y}_{l,N})$ the corresponding result of denoising. Then

$$\mathbf{r} = \frac{1}{L} \sum_{l=1}^{L} \sqrt{\frac{(Y_l - X_l)^2}{(\tilde{Y}_l - X_l)^2}} \tag{5}$$

defines the noise reduction factor which quantifies signal improvement owing to the filter process.

We determined $\mathbf{r}$ for test signals contaminated with white noise, using noise amplitudes ranging from 25% - 150%, and embedding dimensions ranging from 16 - 128 (Figure 2a, Figure 3). Five repetitions for each parameter configuration were calculated using 5 embedded trials each. In the case of $\lambda \leq 2$, the *noise reduction factor* was quite stable against changes of noise levels but depended on embedding dimension m and thresholding coefficient λ. Best performance was achieved for $1.0 \leq \lambda \leq 2.0$ ($\mathbf{r}_{max}^{m=128,\lambda=2.0} = 4.7$). In the case of $\lambda > 4.0$, most signal components were rejected, and as a result, the *noise reduction factor* $\mathbf{r}$ increased linearly with noise levels, as expected. Figure 2b and

Figure 4 depict effects of denoising of 5 test signals contaminated with in-band noise. In comparison to white noise the performance decreased, but neverthe-less, enabled satisfactory denoising for $0.5 \leq \lambda \leq 1.0$ ($\mathbf{r}_{max}^{m=128,\lambda=1.0} = 1.6$). Within this range, the *noise reduction factor* $\mathbf{r}$ depended weakly on noise lev-els. Note that the embedding dimension must be sufficiently high ($m = 128$) to find true neighbors.

In order to simulate EPs/ERPs with several time-varying components, we used 5 test signals which were again a concatenation of different Gaussian functions, each, however, randomly shifted in time (Figure 2c and Figure 5). In contrast to test signals with time fixed components, "optimal" embedding dimension depended on the thresholding coefficient λ. Higher values of λ required lower embedding dimensions and vice versa. Best results were achieved for $0.5 \leq \lambda \leq 2.0$ ($\mathbf{r}_{max}^{m=128,\lambda=1.0} = 3.2$).

Even for high noise levels, the proposed denoising scheme preserved finer struc-tures of original test signals in all simulations. Moreover, the reconstructed sequences were closer to the test signals than the corresponding averages, es-pecially for time varying signals. Power spectra showed that denoising took part in all frequency bands and was quite different from common low-, or band-pass filtering. Simulation indicated that "optimal" values of the thresh-olding coefficient were in the range $0.5 \leq \lambda \leq 2.0$. Best embedding dimension was found to be $m = 128$, since the ongoing background EEG can be assumed to be in-band with ERPs. The filter performance was quite stable against the number of embedded sequences, at least for $L = 5, 10, 20$.

4 Real data

4.1 Data Acquisition

We analyzed event related potentials recorded intracerebrally in patients with pharmacoresistent focal epilepsy [26]. Electroencephalographic signals were recorded from bilateral electrodes implanted along the longitudinal axis of the hippocampus. Each electrode carried 10 cylindrical contacts of nickel-chromium alloy with a length of 2.5 mm and an intercontact distance of 4 mm. Signals were referenced to linked mastoids, amplified with a bandpass filter setting of 0.05 - 85.00 Hz (12dB/oct.) and, after 12 bit A/D conversion, continuously written to a hard disk using a sampling interval of $5760\mu s$. Stimulus related epochs spanning 1480 ms (256 sample points) including a 200 ms pre-stimulus baseline were extracted from recorded data. The mean of the pre-stimulus baseline was used to correct possible amplitude shifts of the following ERP epoch.

In a *visual odd-ball paradigm* 60 rare (letter $< x >$, targets) and 240 frequent stimuli (letter $< o >$, distractors) were randomly presented on a computer monitor once every $1200 \pm 200ms$ (duration: 100 ms, probability of occurrence: 1 ($< x >$) : 5 ($< o >$)). Patients were asked to press a button upon each rare target stimulus. This pseudo-random presentation of rare stimuli in combination with the required response is known to elicit the mesial temporal lobe (MTL) P300 potential in recordings from within the hippocampal formation [27] (cf. Figure 1).

4.2 Results

By simulation, we estimated a range in which "optimal" parameters of the filter can be expected. However, the quality of denoising ERP sequences could

not be estimated, because the clean signal was not known a priori. A rough estimation of filter performance was only possible by a comparison to ERP averages. Taking into account results of simulation as well as ERP averages, we estimated $\lambda = 0.6$ and $m = 128$ to be the best configuration.

Based on the empirical fact that specific ERP components exhibit peak amplitudes within a narrow time range related to stimulus onset, we defined a maximum allowed time jitter of ± 20 sample points ($\approx 116ms$) true neighbors are assumed to be restricted to. This accelerated the calculation time and avoided false nearest neighbors. Figure 6 depicts several ERPs recorded from different electrode contacts within the hippocampal formation. The number of embedded sequences was chosen as $L = 8$. Comparing averages, we can expect that the filter extracted the most relevant MTL-P300 components. Even for low amplitude signals reconstruction was possible, exhibiting higher amplitudes in single trial data than in averages. As corresponding power spectra show, the 50 Hz power line was reduced but not eliminated after filtering. Especially low amplitude signals showed artifacts based on the 50 Hz power line.

5 Conclusion

In this study, we introduced a new wavelet based method for nonlinear noise reduction of single trial EPs/ERPs. We employed advantages of methods developed for the paradigm of deterministic chaotic systems, that allowed denoising of short and time variant EP/ERP sequences without assuming fully synchronized or stationary EEG.

Denoising via wavelet shrinkage does not require a priori assumptions about constrained dimensions, as is usually required for other techniques (e.g., singu-

lar value decomposition). Besides, it is more straight forward using thresholds depending on means and variances rather than initial assumptions about constrained embedding dimensions. Moreover, the local calculation of thresholds in state space enables focusing on specific frequency scales, which may be advantageous in order to extract signal components located within both narrow frequency bands and narrow time windows.

Extension of our denoising scheme to other types of signals seems to be possible, however, demands further investigations, since "optimal" filter parameters highly depend on signal characteristics. In addition, the *noise reduction factor* **r** does not consider all imaginable features of signals investigators are possibly interested in, so that other measures may be more advantageous in specific cases.

So far, we have not considered effects of smoothing the edges of signal sequences. But since delay vectors as well as corresponding wavelet coefficients hold information locally, we can assume artifacts to be also constrained to the edges which we were not interested in.

In conclusion, the proposed denoising scheme represents a powerful noise reduction technique for transient signals of short duration, like ERPs.

References

[1] F. H. Lopes da Silva, A. S. Givens, and A. Remond, editors. *Handbook of Electroencephalography and Clinical Neurophysiology.* Elsevier Science Publisher B.V., Amsterdam, 1986.

[2] S. Cerutti, G. Basselli, and G. Pavesi. Single sweep analysis of visual evoked potentials through a model of parametric identification. *Biol Cybern*, 56:111, 1987.

[3] H. J. Heinze and H. Künkel. ARMA-filtering of evoked potentials. *Meth. Inform. Med.*, 23:29, 1984.

[4] C. D. Woody. Characterisation of an adaptive filter for the analysis of variable latency neuroelectric signals. *Med Biol Eng*, 5:539, 1967.

[5] D. O. Walter. A posteriori Wiener filtering of average evoked response. *Electroencephalogr Clin Neurophysiol (Suppl)*, 27:61, 1969.

[6] D. J. Doyle. Some comments on the use of Wiener filtering in the estimation of evoked potentials. *Electroencephalogr Clin Neurophysiol*, 38:533, 1975.

[7] J. P. De Weerd. A posteriori time-varying filtering of averaged evoked potentials. I. introduction and conceptual basis. *Biol Cybern*, 41:211, 1981.

[8] J. P. De Weerd and J. I. Kap. A posteriori time-varying filtering of averaged evoked potentials. II. mathematical and computational aspects. *Biol Cybern*, 41:223, 1981.

[9] C. S. Burrus, R. A. Copinath, and H. Guo. *Wavelets and wavelet transforms.* Prentice Hall, New Jersey, 1998.

[10] C. K. Chui. *Introduction to wavelets.* Academic Press, San Diego, 1992.

[11] I. Daubechies. *Ten lectures on wavelets.* Society for Industrial and Applied Mathematics, Pennsylvania, 1992.

[12] R. R. Coifman and M. Y. Wickerhauser. Wavelets, adapted waveforms and de-noising. *Electroencephalogr Clin Neurophysiol (Suppl.)*, 45:57, 1996.

[13] V. J. Samar, K. P. Swartz, and M. R. Raghuveer. Multiresolution analysis of event-related potentials by wavelet decomposition. *Brain Cogn*, 27:398, 1995.

[14] S. J. Schiff, A. Aldrouby, M. Unser, and S. Sato. Fast wavelet transform of EEG. *Electroencephalogr Clin Neurophysiol*, 91:442, 1994.

[15] S. J. Schiff, J. Milton, J. J. Heller, and S. Weinstein. Wavelet transforms and surrogate data for electroencephalographic spike and seizure detection. *Opt Eng*, 33:2162, 1994.

[16] D. L. Donoho, I. M. Johnstone, and B. W. Silverman. De-noising by soft-thresholding. *IEEE Trans Inform Theor*, 41:613, 1995.

[17] O. Bertrand, J. Bohorquez, and J. Pernie. Time frequency digital filtering based on an invertible wavelet transform: An application to evoked potentials. *IEEE Trans Biomed Eng*, 41:77, 1994.

[18] E. A. Bartink, K. J. Blinowska, and P. J. Durka. Single evoked potential reconstruction by means of wavelet transform. *Biol Cybern*, 67:175, 1992.

[19] E. A. Bartink, K. J. Blinowska, and P. J. Durka. Wavelets: New method of evoked potential analysis. *Med Biol Eng Comput*, 30:125, 1992.

[20] H. Kantz and T. Schreiber. *Nonlinear time series analysis*. Cambridge University Press, Cambridge, 1997.

[21] T. Schreiber and D. T. Kaplan. Nonlinear noise reduction for electrocardiograms. *Chaos*, 6:87, 1995.

[22] P. Grassberger, R. Hegger, H. Kantz, C. Schaffrath, and T. Schreiber. On noise reduction methods for chaotic data. *Chaos*, 41:127, 1993.

[23] S. G. Mallat. Multiresolution approximation and wavelet orthonormal bases of L^2. *IEEE Trans Am Math Soc*, 315:69, 1989.

[24] S. G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans Patt Recog Mach Intel*, 11:674, 1989.

[25] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer. Testing of nonlinearity in time series: The method of surrogate data. *Physica D*, 58:77, 1992.

[26] T. Grunwald, H. Beck, K. Lehnertz, I. Blümcke, N. Pezer, M. Kutas, M. Kurthen, H. M. Karakas, D. Van Roost, O. D. Wiestler, and C. E. Elger. Limbic P300s in temporal lobe epilepsy with and without Ammon's horn sclerosis. *Eur J Neurosci*, 11:1899, 1999.

[27] A. Puce, R. M. Berkovic, G A. Donnan, and P. F. Baldin. Limbic P3 potentials, seizure localization, and surgical pathology in temporal lobe epilepsy. *Annals of Neurology*, 26:377, 1989.

Figure captions:

Fig. 1: Examples of averaged ERPs recorded along the longitudinal axis of the hippocampal formation in a patient with epilepsy. Randomized presentation of *target* and *standard* stimuli is known to elicit the mesial temporal lobe P300, a negative deflection peaking at about 500 ms after stimulus onset (cf. Sect. 4.1 for more details). Letters (a), (b), and (c) indicate recordings used for single trial analysis (cf. Figure 6).
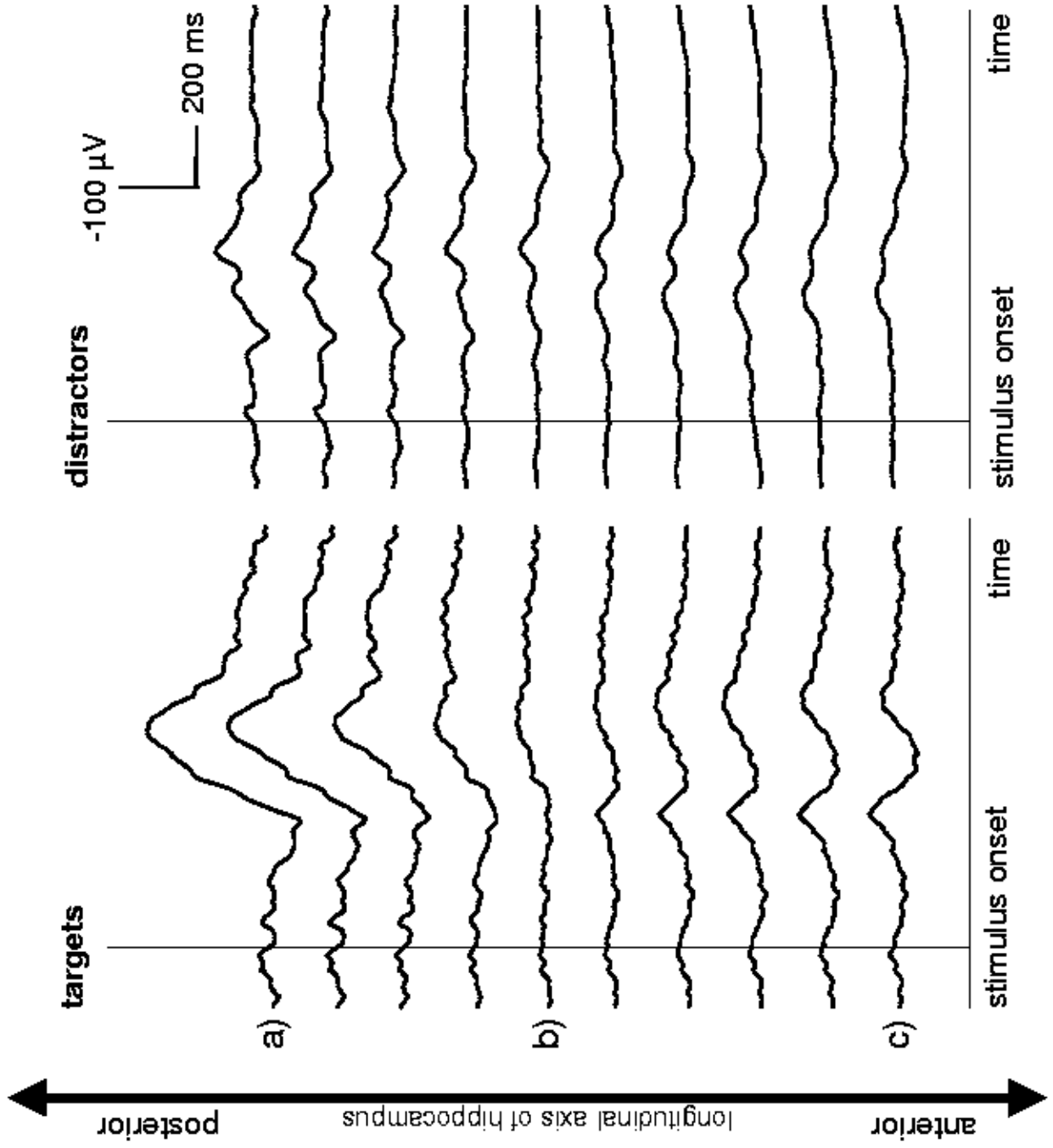
Fig. 2: Results of denoising test signals. Parts a) and b): contamination with white noise and in-band noise. Part c): time varying signal components and white noise contamination (see text for more details). Five calculations for each parameter configuration have been executed to determine standard deviations.
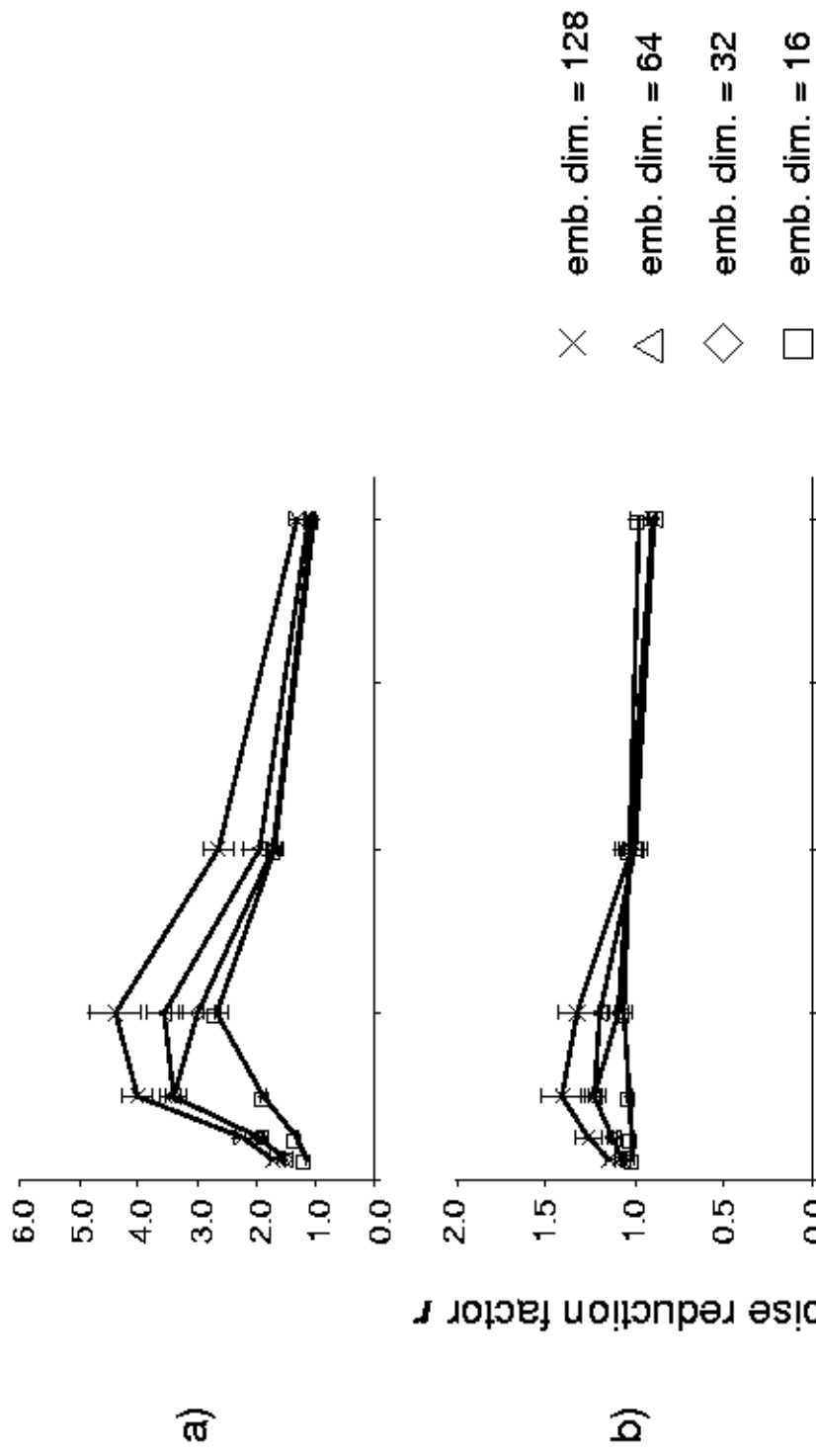
Fig. 3: Nonlinear denoising applied to white noise contaminated test signals (5 sequences embedded, each 256 sample points, randomly shifted in time (std. dev.: 20 sample points, max. shift: 40 sample points), noise amplitude 75%, $m = 128$, $\tau = 1$, $\lambda = 1.5$). Power spectra in arbitrary units. For state space plots we used a time delay of 25 sample points.

Fig. 4: Same as Figure 3 but for in-band noise and $\lambda = 0.75$.

Fig. 5: Same as Figure 3 but for Gaussian functions each randomly shifted in time and $\lambda = 0.75$.

Fig. 6: Examples of denoised MTL-P300 potentials (cf. Figure 1). Power spectra in arbitrary units. For state space plots we used a time delay of 25 sample points.

noise reduction factor r

thresholding coefficient λ

emb. dim. = 128

emb. dim. = 64

emb. dim. = 32

emb. dim. = 16

a)

b)

c)

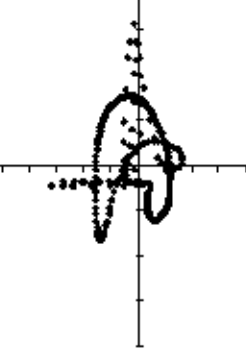| example:	1	2	3	4	5	average	power spectrum	state space
test signal								
test signal + white noise								
denoised test signal								
residual noise								

example: | 1 | 2 | 3 | 4 | 5 | average | power spectrum | state space
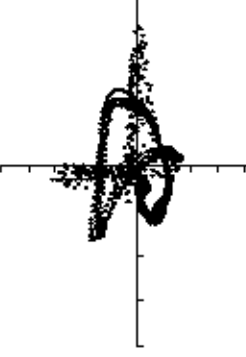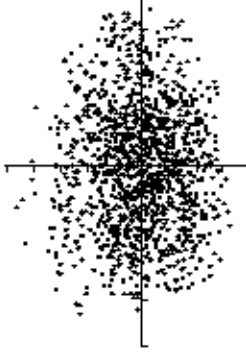
test signal

test signal + in-band noise

denoised test signal

residual noise

example:	1	2	3	4	5	average	power spectrum	state space
test signal								
test signal + white noise								
denoised test signal								
residual noise								

example:

state space

power spectra

averages

a) posterior

raw
denoised
residual

b) longitudinal axis of hippocampus

raw
denoised
residual

c) anterior

raw
denoised
residual

1 2 3 4 5